

The background of the slide features the Archive-IT logo, which consists of a large, light blue triangle containing a white stylized 'A' with a blue circle in the center. Below the triangle, the text 'ARCHIVE-IT' is written in a bold, white, sans-serif font.

So, you've inherited an Archive-It account...

Part 2: Introduction to seeds

ARCHIVE-IT

Introduction

Facilitators:

- Karl Blumenthal, Senior Web Archivist
- Sarah Beth Seymore, Program Officer



Learning objectives:

1. How to **select, configure, organize**, and **access** seeds
2. When and how to find help

Download and reference:

→ <https://archive.org/details/AIT-seed-training>

Introduction

Facilitators:

- Karl Blumenthal, Senior Web Archivist
- Sarah Beth Seymore, Program Officer



Learning objectives:

1. How to **select, configure, organize**, and **access** seeds
2. When and how to find help

Download and reference:

Follow the links to our help docs

→ <https://archive.org/details/AIT-seed-training>



Quick reference

1. [How to select and add seed URLs](#)

A summary of seed considerations and examples.

2. [Scoping recommendations for specific sites](#)

Rules and configurations known to help archive common kinds of seeds.

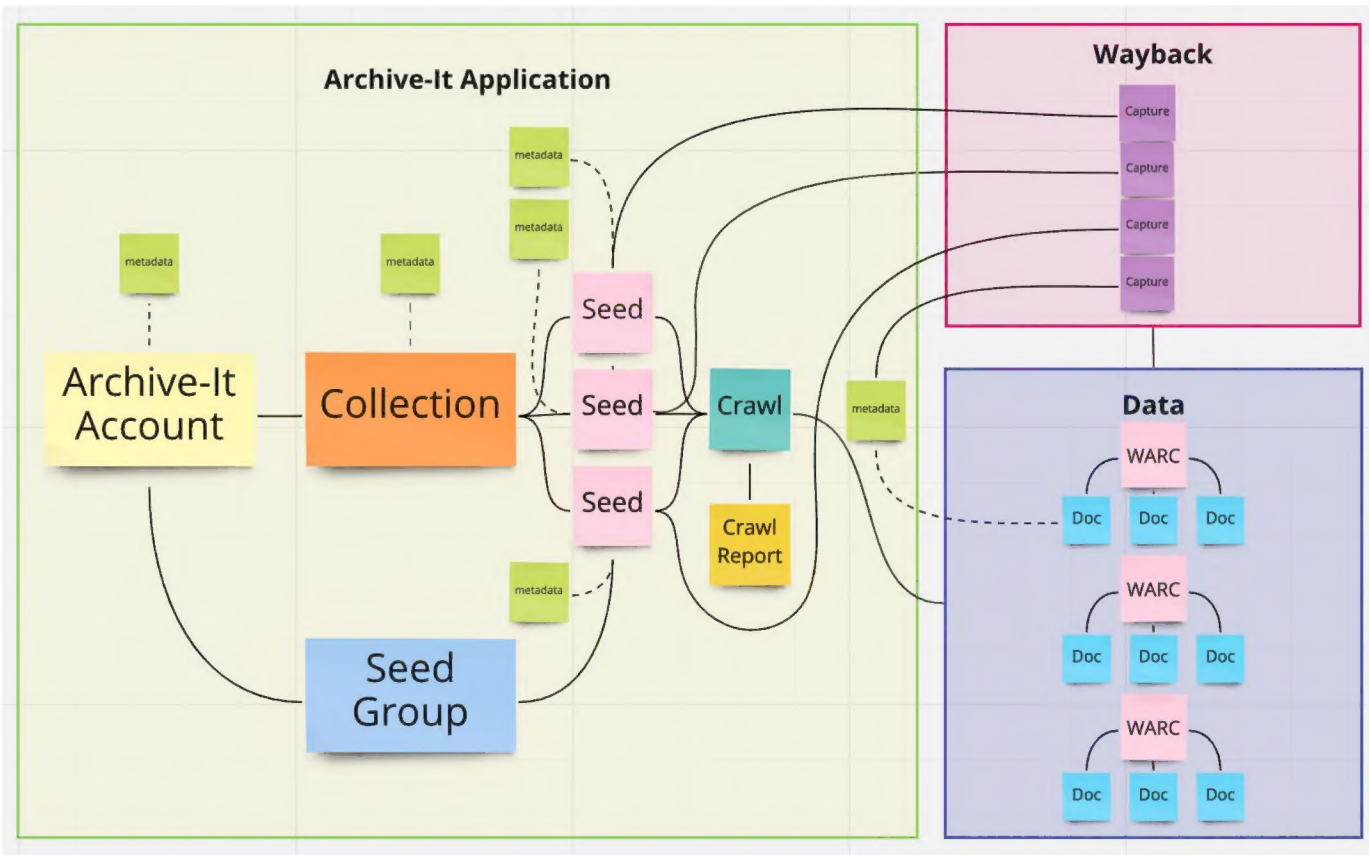
3. [Archive-It Office Hours](#)

Schedule a time to answer your seed questions and explore options with the Internet Archive's web archivists.

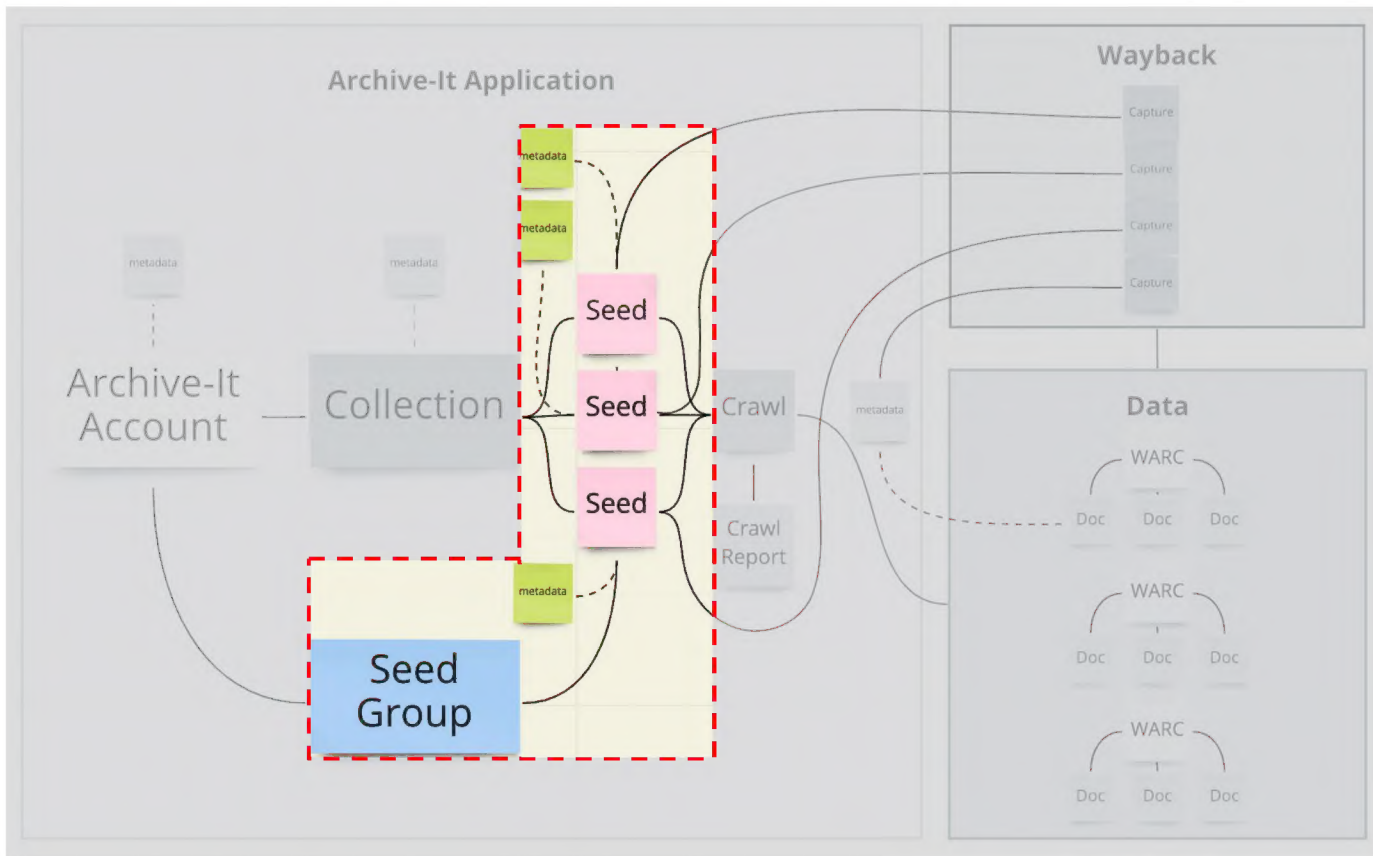
4. [Community Help](#)

Ask us and your colleagues about selection policies, descriptive standards, management workflows, and more.

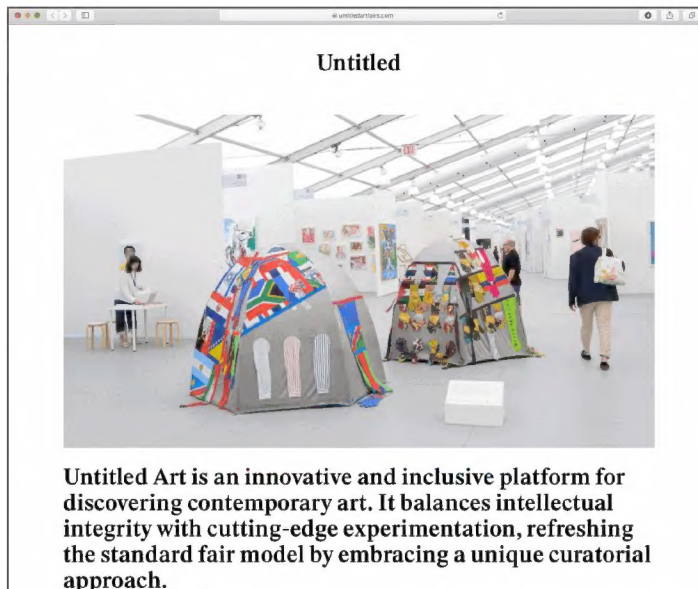
What is a seed



What is a seed



Selecting seeds

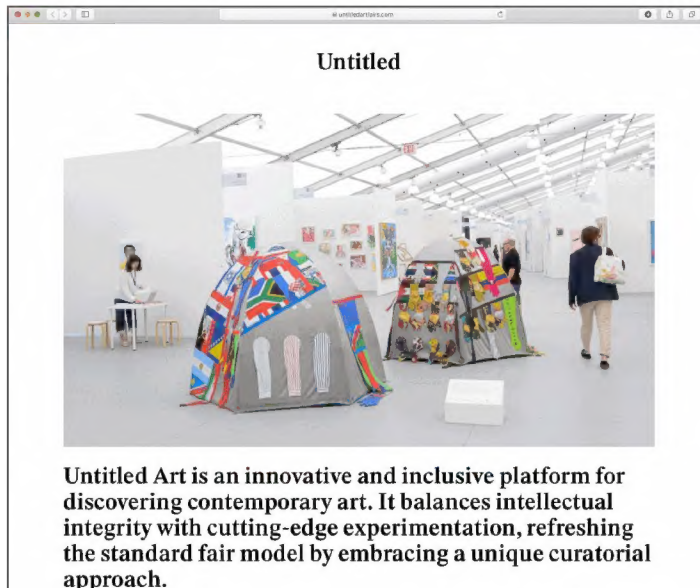


When selecting new or evaluating existing seeds in your web archive collections, ask yourself:

- ★ **Where is it?**
- ★ **What is it?**
- ★ **What are its restrictions?**

Sample selections for the Collaborative ART Archive (CARTA), <https://carta.archive-it.org/>

Seed URLs



Sample selections for the Collaborative ART Archive (CARTA), <https://carta.archive-it.org/>

Each seed URL can represent:

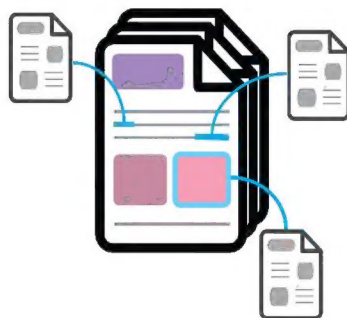
1. **A whole web domain**
<https://archive.org/>
2. **A host or directory**
<https://mastodon.archive.org/@internetarchive/>
3. **A single page or document**
<http://archive.org/services/collection-rss.php>

Seed types

Standard



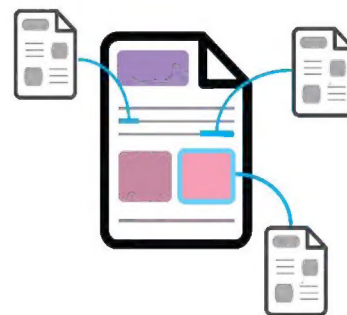
Standard +



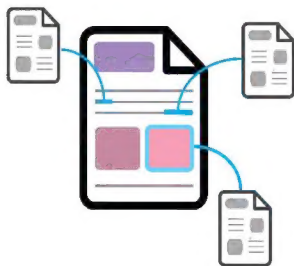
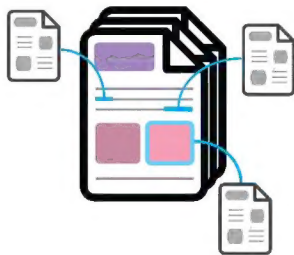
One Page



One Page +



Seed scope



Configure seeds optionally to:

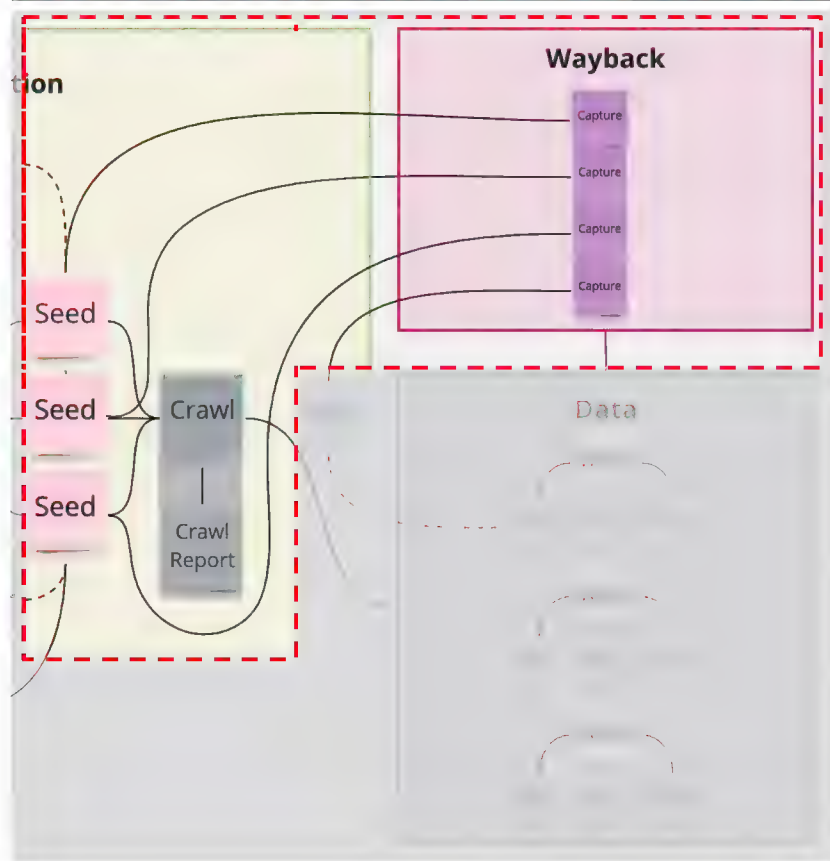
- Limit data
- Ignore blocks and delays
- Exclude audio and video
- Accept or exclude document if...

Give crawling a try - test are free!

Best practices for a first-time crawler:

1. Seeds: Start with ≤ 10 for an initial test if possible
 2. Seed type: **Standard** or **One Page** - avoid “plus” types for now
 3. Crawl type: **Test**
 4. Crawl limit: **3 days**, skip data and document limits for now
 5. Crawling technology: **Brozzler**, for high fidelity collecting
- [Enable email notifications](#) to know when your crawls complete.

Access controls



Seeds can be:

1. **Public** - Listed on Archive-It.org
2. **Private** - Unlisted
- ★ **Shared** - Listed in collection/s other than the original

Seed “Groups” further organize seeds for faceted browsing.

Describing seeds

Title: Acoso Online

URL: <https://acoso.online/>

Description: Proyecto nacido en el 2017 como una respuesta urgente para las mujeres y personas de la comunidad de la diversidad sexogenérica que día a día sufren violencia de género en línea y difusión de contenido íntimo sin consentimiento. Han trabajado también en contenidos más amplios de violencia de género online que puedes encontrar en el sitio web. El sitio brinda, principalmente, orientación a víctimas de 19 países de América Latina y El Caribe, además de España. Contiene orientación legal, judicial y comunitaria, además de qué hacer en las plataformas más populares cuando se enfrenta un caso.

Captured once on December 07, 2022

Language: español, inglés, portugués

Coverage: Regional

Temas: Acoso y hostigamiento sexual, Tecnología feminista

Title: Centro de Mujeres Afrocostarricenses

URL: <https://afrodescendientes.org/>

Description: Iniciativa política de mujeres afrocostarricenses para trabajar con temas relacionados a género, participación política y raza. Nace en 1992 y está conformada por dirigentes, profesionales y mujeres de base en diversos campos, encargadas de incentivar iniciativas de desarrollo justas, que busquen eliminar todas las formas de discriminación, especialmente la Discriminación Étnica y de Género.

Captured once on February 20, 2023

Language: español

Coverage: Costa Rica

Temas: Antirracismo, Derechos políticos

References and case studies:

- [Descriptive Metadata for Web Archiving](#), OCLC.
- [Metadata Application Profile for Description of Websites with Archived Versions](#), NYARC.
- [Describing Archived Websites](#), UNC Chapel Hill.

Activismos feministas en América Latina,
<https://idrhuu.org/huellasincomodas/webarchive>



What am I collecting? And why?

- ❖ Get to know your seeds
 - Download an inventory of all seeds in your account, including administrative and descriptive metadata, anytime here:
<https://partner.archive-it.org/api/seed?format=csv> (example)
- ❖ Browse your colleagues' collecting policies
 - [DC Public Library](#)  | [New Brunswick Free Public Library](#) 
 - [Austin Seminary](#)  | [Kansas Archive-It Consortium](#) 
 - UC San Diego - [Policy](#)  | [Libguide](#) 

Help when you need it

Q: Why is my seed URL reporting an error or a “not crawled” status?

↳ [Chat live](#) (9AM - 12PM Pacific, Tues-Thurs) or [submit a ticket](#) to the Archive-It team.

Help when you need it

Q: Why is my seed URL reporting an error or a “not crawled” status?

↳ Chat live (9AM - 12PM Pacific, Tues-Thurs) or submit a ticket to the Archive-It team.

Q: How can I make my seeds collect more of what I need (and less of what I don't)?

↳ [Book office hours](#) to dig in deeper with a web archivist.

Help when you need it

Q: Why is my seed URL reporting an error or a “not crawled” status?

↳ Chat live (9AM - 12PM Pacific, Tues-Thurs) or submit a ticket to the Archive-It team.

Q: How can I make my seeds collect more of what I need (and less of what I don't)?

↳ Book office hours to dig in deeper with a web archivist.

Q: How can I reorganize or export seeds across my account for access?

↳ [Schedule a custom training](#) for you or your team.

Help when you need it

Q: Why is my seed URL reporting an error or a “not crawled” status?

↳ Chat live (9AM - 12PM Pacific, Tues-Thurs) or submit a ticket to the Archive-It team.

Q: How can I make my seeds collect more of what I need (and less of what I don't)?

↳ Book office hours to dig in deeper with a web archivist.

Q: How can I reorganize or export seeds across my account for access?

↳ Schedule a custom training for you or your team.

Q: How are my colleagues selecting their seed URLs?

↳ Read, post, and share posts in the [Archive-It partners' community forum](#).

Coming up next



Crawl space: An Archive-It community forum

- July 8, 2026, at 11AM Pacific, 2PM Eastern
- “How not to start a web archiving program in two weeks”



Feature release & training: Custom crawl scheduling

- July 22, 2026, at 11AM Pacific, 2PM Eastern
- Learn the new and improved options for scheduling recurring crawls



So you've inherited an Archive-It account, Part 3: Troubleshooting basics

- August 26, 2026, at 11AM Pacific, 2PM Eastern
- Crawl review, quality assurance, and help when you need it